PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

Assistant Commissioner for Patents
United States Patent and Trademark Office
Box PCT
Washington, D.C.20231
ÉTATS-UNIS D'AMÉRIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 10 September 1999 (10.09.99)	
International application No. PCT/SG98/00009	**Applicant's or agent's file reference** SGS/50567
International filing date (day/month/year) 12 February 1998 (12.02.98)	**Priority date** (day/month/year)
Applicant ABSAR, Mohammed, Javed et al	

1. The designated Office is hereby notified of its election made:

   [X] in the demand filed with the International Preliminary Examining Authority on:

   19 August 1999 (19.08.99)

   [ ] in a notice effecting later election filed with the International Bureau on:

2. The election [X] was [ ] was not made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO **34, chemin des Colombettes** **1211 Geneva 20, Switzerland** Facsimile No.: (41-22) 740.14.35	Authorized officer C. Carrié Telephone No.: (41-22) 338.83.38

ENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference	FOR FURTHER ACTION	see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.
SGS/50567		
International application No.	International filing date *(day/month/year)*	(Earliest) Priority Date *(day/month/year)*
PCT/SG 98/00009	12/02/1998	
Applicant		
SGS-THOMSON MICROELELECTRONICS ASIA PACIFIC et al.		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 2 sheets.

[X] It is also accompanied by a copy of each prior art document cited in this report.

1. [ ] **Certain claims were found unsearchable** (see Box I).

2. [ ] **Unity of invention is lacking** (see Box II).

3. [ ] The international application contains disclosure of a **nucleotide and/or amino acid sequence listing** and the international search was carried out on the basis of the sequence listing
   - [ ] filed with the international application.
   - [ ] furnished by the applicant separately from the international application,
     - [ ] but not accompanied by a statement to the effect that it did not include matter going beyond the disclosure in the international application as filed.
   - [ ] Transcribed by this Authority

4. With regard to the **title**,
   - [X] the text is approved as submitted by the applicant
   - [ ] the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,
   - [X] the text is approved as submitted by the applicant
   - [ ] the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this International Search Report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is:

   Figure No. 2
   - [ ] as suggested by the applicant.
   - [X] because the applicant failed to suggest a figure.
   - [ ] because this figure better characterizes the invention.

   [ ] None of the figures.

INTERNATIONAL SEARCH REPORT

Int[illegible]al Application No
P[illegible]G 98/00009

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 H04B1/66

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 H04B H03M

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MCKINNEY: "Digital Audio Compression Standard AC-3" 20 December 1995 , ADVANCED TELEVISION SYSTEMS COMITTEE XP002075746 cited in the application see paragraph 7.1.1 see paragraph 7.1.2	1,3-6, 17,19
A	see paragraph 8.2.8 -----	31

☐ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 October 1998	19/10/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Bossen, M

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference SGS/50567	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. PCT/SG98/00009	International filing date *(day/month/year)* 12/02/1998	Priority date *(day/month/year)* 12/02/1998
International Patent Classification (IPC) or national classification and IPC H04B1/66		
Applicant SGS-THOMSON MICROELELECTRONICS ASIA PACIFIC et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 6 sheets, including this cover sheet.

   ☒ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of 12 sheets.

3. This report contains indications relating to the following items:

   I ☒ Basis of the report
   II ☐ Priority
   III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   IV ☐ Lack of unity of invention
   V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   VI ☐ Certain documents cited
   VII ☒ Certain defects in the international application
   VIII ☒ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
19/08/1999	23.05.2000
Name and mailing address of the international preliminary examining authority: European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer Ciccarese, C Telephone No. +49 89 2399 7302

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/SG98/00009

I. Basis of the report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

Description, pages:

1,2,4,7-20	as originally filed	
3,5,6	with telefax of	07/04/2000

Claims, No.:

1-37	with telefax of	07/04/2000

Drawings, sheets:

1/4-4/4	as originally filed	

2. The amendments have resulted in the cancellation of:

☐ the description, pages:
☐ the claims, Nos.:
☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

Re Item V

Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. D1: MCKINNEY: 'Digital Audio Compression Standard AC-3' 20 December 1995, ADVANCED TELEVISION SYSTEMS COMMITTEE XP002075746 cited in the application

2. The application belongs to the field of digital signal processing: samples of an audio signal are sent in frames, subdivided in blocks or sets. The samples are represented in mantissa-exponent pairs, and the goal is to code the exponents in an effective way.

3. Claims 1, 16 and 29 involve an inventive step: the method proposed in D1 does not foresee the use of different differential coding limits for different exponent sets, which results in a more effective coding, nor would this improvement have been obvious for the skilled person.

4. The dependent claims concern advantageous embodiments of the subject-matter of the independent claims and thus their subject-matter is also considered to be novel and inventive.

Re Item VII

Certain defects in the international application

1 The features of the claims are not provided with reference signs placed in parentheses (Rule 6.2(b) PCT).

2 The independent claims are not in the two-part form in accordance with Rule 6.3(b) PCT, which in the present case would be appropriate, with those features known in combination from the prior art (document D1) being placed in a preamble (Rule 6.3(b)(i) PCT) and with the remaining features being included in a characterising portion (Rule 6.3(b)(ii) PCT).

Re Item VIII
Certain observations on the international application

1 All the features of claim 16 are present in claim 1.
The aforementioned claims therefore lack conciseness. Moreover, lack of clarity of the claims as a whole arises, since the plurality of independent claims makes it difficult, if not impossible, to determine the matter for which protection is sought, and places an undue burden on others seeking to establish the extent of the protection.

Hence, claims 1 and 16 do not meet the requirements of Article 6 PCT.

2 Claims 13, 26 and 35 are not clear, since it is not clear what the quantity 'i' represents.

3 It seems that in claims 12, 25 and 34 and in the corresponding passages of the description the wording

"$f(y_i)$ is +1 if $y \geq 0$..."

should read

"$f(y_i)$ is +1 if $y_i \geq 0$...".

4 The text at the beginning of the last line of pages 3 and 5 is partially unreadable.

5 The text at the beginning of page 3 does not match that at the end of page 2.

6 The text at the beginning of page 4 does not match that at the end of page 3.

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims	1-37
	No:	Claims	
Inventive step (IS)	Yes:	Claims	1-37
	No:	Claims	
Industrial applicability (IA)	Yes:	Claims	1-37
	No:	Claims	

2. Citations and explanations

see separate sheet

VII. Certain defects in the international application

The following defects in the form or contents of the international application have been noted:

see separate sheet

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

see separate sheet

09/622147
528 Rec'd PCT/PTO 11 AUG 2000

computational resources.

Summary of the Invention

In accordance with the present invention, there is provided a method for processing data in an audio data encoder, the data comprising a sequence of exponent sets each comprising a plurality of exponents, comprising the steps of:

determining a first variation of exponent values within a first exponent set;

determining a second variation of exponent values between said first exponent set and each subsequent exponent set in said sequence; and

assigning an exponent coding strategy to the first exponent set based on the determined first and second variations, wherein the exponent coding strategy is assigned from a plurality of exponent coding strategies having different differential coding limits.

In one form of the invention, the method includes selecting one of said subsequent exponent sets on the basis of said first variation and assigning an exponent re-use coding strategy to the selected exponent set and any exponent sets in said sequence between the first exponent set and the selected exponent set.

Preferably the method includes a step of determining a second variation between consecutive exponents in said first exponent set, wherein the exponent coding strategy for said first exponent set is selected on the basis of said first and second variations.

The step of selecting the exponent coding strategy for the first exponent set is preferably performed using neural network processing. In that case, the neural network processing may include a feature extraction stage in which said sequence of exponent sets is processed to determine said first variation values, a weighted routing stage in which said first variation values are weighted according to predetermined weighting values and routed to inputs of a first neural layer, a selection stage in which an output of the first neural layer is selected, and an output processing stage in which a coding strategy is assigned to said first exponent set t of said selection stage and said second variation.

$$\Gamma\begin{bmatrix}\gamma_0\\ \gamma_1\\ \gamma_2\\ \cdots\\ \cdots\\ \gamma_{b-1}\end{bmatrix} = \begin{bmatrix}f(\gamma_0)\\ f(\gamma_1)\\ f(\gamma_2)\\ \cdots\\ \cdots\\ f(\gamma_{b-1})\end{bmatrix}$$

where $f(\gamma_i)$ is $+1$ if $\gamma \geq 0$ else it is 0,
Y represents outputs of the first neural layer,
T are threshold values determined during a training phase,
w are weighting values determined during the training phase, and
b is the number of exponent sets in the sequence comprising said data.

In one form of the invention, the selection stage comprises selecting an output y_a of the first neural layer such that $y_a = 1$ and a is maximum for $i<a<b$. The plurality of exponent coding strategies may comprise strategies S_1, S_2, ... , S_c, where $c \leq b$, corresponding to respective differential coding limits 1, 2, ... ,c, wherein the exponent coding strategy S_Y assigned to said first exponent set E_i is selected according to

$$\gamma = \max[\min(a+1, \sigma(E_i)), 1]$$

where $\sigma(E_i) = floor((\sum_j \| e_{i,j+i} - e_{i,j} \| / n) + 0.5)$.

The exponent sets in the sequence between $i+1$ and a, inclusive, can then be assigned the exponent re-use coding strategy.

The present invention also provides a method for coding audio data comprising a sequence of exponent sets each comprising a plurality of exponents, comprising the steps of:

determining a first variation of exponent values between a first exponent set in the sequence and each subsequent exponent set in said sequence;

selecting an exponent coding strategy for said first exponent set from a plurality of rategies on the basis of said first variation; and

coding said first exponent set according to the selected exponent coding strategy, wherein each of the plurality of exponent coding strategies corresponds to a different differential coding limit.

The invention also provides a digital audio encoder in which audio data is transformed into coefficients having mantissas and exponents arranged in a sequence of sets, having:

a first variation processor coupled to receive the exponents of sets from said sequence and determine a first variation of exponent values between a first set and a plurality of subsequent sets in the sequence;

a second variation processor coupled to receive the exponents of said first set and determine a second variation between consecutive exponent values within said first set; and

a neural network processor coupled to receive said first and second variations and select and assign an exponent coding strategy to said first set from a plurality of coding strategies on the basis of said first and second variations, wherein each of the plurality of coding strategies correspond to a different differential coding limit.

Brief Description of the Drawings

The invention is described in greater detail hereinafter, by way of example only, with reference to several embodiments, and having regard to the accompanying drawings, wherein:

Figure 1 is a diagrammatic illustration of the data structure of an encoded audio data bit stream, having frames and blocks;

Figure 2 is a functional block diagram of an audio encoder;

Figure 3 is a diagrammatic illustration of mapping of exponent sets to coding strategies; and

Figure 4 is a diagrammatic illustration of a neural network system for selection of exponent coding strategies.

Detailed Description of the Preferred Embodiments

Claims:

1. A method for processing data in an audio data encoder, the data comprising a sequence of exponent sets each comprising a plurality of exponents, comprising the steps of:

determining a first variation of exponent values within a first exponent set;

determining a second variation of exponent values between said first exponent set and each subsequent exponent set in said sequence; and

assigning an exponent coding strategy to the first exponent set based on the determined first and second variations, wherein the exponent coding strategy is assigned from a plurality of exponent coding strategies having different differential coding limits.

2. A method as claimed in claim 1, including a step of coding said first exponent set according to the assigned exponent coding strategy.

3. A method as claimed in claim 2, including a step of assigning an exponent coding strategy to at least one subsequent exponent set based on the corresponding determined second variation.

4. A method as claimed in claim 3, wherein the plurality of exponent coding strategies includes an exponent set re-use strategy which is assigned to the at least one subsequent exponent set.

5. A method as claimed in claim 4, including a step of coding said first exponent set and said at least one subsequent exponent set according to the corresponding assigned coding strategies.

6. A method as claimed in claim 1 or 5, wherein the steps of determining the first and second variations are performed utilising neural network processing.

7. A method as claimed in claim 6, wherein the neural network processing includes first

and second neural layers.

8. A method as claimed in claim 6, wherein the neural network processing comprises a feature extraction stage in which said sequence of exponent sets is utilised to determine said second variations, a weighted routing stage in which said second variations are weighted according to predetermined weighting values and routed to inputs of a first neural layer, a selection stage in which an output of the first neural layer is selected, and an output processing stage in which a coding strategy is assigned to said first exponent set based on said first variation and the output of said selection stage.

9. A method as claimed in claim 8, wherein a coding strategy is assigned to at least one subsequent exponent set on the basis of the output of said selection stage.

10. A method as claimed in claim 9, wherein the coding strategy assigned to the at least one subsequent exponent set is an exponent re-use strategy.

11. A method as claimed in claim 8, wherein the feature extraction stage comprises determining

$$\mathrm{Adiff}(E_i, E_j) = (\Sigma_m \, |e_{i,m} - e_{j,m}|)/n$$

where Adiff is said second variation,

E_i is said first exponent set and E_j is a subsequent exponent set with $j > i$,

$E_i = (e_{i,0}, e_{i,1}, e_{i,2}, \ldots, e_{i,n-1})$,

$E_j = (e_{j,0}, e_{j,1}, e_{j,2}, \ldots, e_{j,n-1})$,

n is an integer representing the number of exponents in a said set of exponents,

$m = 0, 1, 2, \ldots, n-1$.

12. A method as claimed in claim 11, wherein the processing carried out by the weighted routing stage and first neural layer includes determining

$$Y = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \cdots \\ \cdots \\ y_{b-1} \end{bmatrix} = \Gamma\left(\begin{bmatrix} 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & w_{1,1} & 0 & \cdots & \cdots & 0 \\ 0 & w_{2,1} & w_{2,2} & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & w_{b-1,1} & w_{b-1,2} & \cdots & \cdots & w_{b-1,b-1} \end{bmatrix} \begin{bmatrix} 0 \\ z_1 \\ z_2 \\ \cdots \\ \cdots \\ z_{b-1} \end{bmatrix} - \begin{bmatrix} 0 \\ T_1 \\ T_2 \\ \cdots \\ \cdots \\ T_{b-1} \end{bmatrix} \right)$$

the operator $\Gamma[\cdot]$ is defined as:

$$\Gamma \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \cdots \\ \cdots \\ \gamma_{b-1} \end{bmatrix} = \begin{bmatrix} f(\gamma_0) \\ f(\gamma_1) \\ f(\gamma_2) \\ \cdots \\ \cdots \\ f(\gamma_{b-1}) \end{bmatrix}$$

where $f(\gamma_i)$ is $+1$ if $\gamma \geq 0$ else it is 0,

Y represents outputs of the first neural layer,

T are threshold values determined during a training phase,

w are weighting values determined during the training phase, and

b is the number of exponent sets in the sequence comprising said data.

13. A method as claimed in claim 12, wherein the selection stage comprises selecting an output y_a of the first neural layer such that $y_a = 1$ and a is maximum for $i < a < b$.

14. A method as claimed in claim 13, wherein the plurality of exponent coding strategies comprises strategies $S_1, S_2, \ldots, S_c$, where $c \leq b$, corresponding to differential coding limits $1, 2, \ldots, c$.

15. A method as claimed in claim 14, wherein the exponent coding strategy S_γ assigned to said first exponent set E_i is selected according to

$$\gamma = \max[\min(a+1, \sigma(E_i)), 1]$$

where $$\sigma(E_i) = \mathrm{floor}((\textstyle\sum_j \| e_{i,j+1} - e_{i,j} \| / n) + 0.5).$$

16. A method for coding audio data comprising a sequence of exponent sets each comprising a plurality of exponents, comprising the steps of:

determining a first variation of exponent values between a first exponent set in the sequence and each subsequent exponent set in said sequence;

selecting an exponent coding strategy for said first exponent set from a plurality of exponent coding strategies on the basis of said first variation; and

coding said first exponent set according to the selected exponent coding strategy, wherein each of the plurality of exponent coding strategies corresponds to a different differential coding limit.

17. A method as claimed in claim 16, including selecting one of said subsequent exponent sets on the basis of said first variation and assigning an exponent re-use coding strategy to the selected exponent set and any exponent sets in said sequence between the first exponent set and the selected exponent set.

18. A method as claimed in claim 16 or 17, including a step of determining a second variation between consecutive exponents in said first exponent set, wherein the exponent coding strategy for said first exponent set is selected on the basis of said first and second variations.

19. A method as claimed in claim 16, wherein the step of selecting the exponent coding strategy for said first exponent set is performed utilising neural network processing.

20. A method as claimed in claim 18, wherein the step of selecting the exponent coding strategy for said first exponent set is performed utilising neural network processing.

21. A method as claimed in claim 20, wherein the neural network processing comprises a feature extraction stage in which said sequence of exponent sets is processed to determine said first variation values, a weighted routing stage in which said first variation values are weighted according to predetermined weighting values and routed to inputs of a first neural layer, a selection stage in which an output of the first neural layer is selected, and an output processing stage in which a coding strategy is assigned to said first exponent set based on the output of said selection stage and said second variation.

22. A method as claimed in claim 21, wherein a coding strategy is assigned to at least one subsequent exponent set on the basis of the output of said selection stage.

23. A method as claimed in claim 22, wherein the coding strategy assigned to the at least one subsequent exponent set is an exponent re-use strategy.

24. A method as claimed in claim 21, wherein the feature extraction stage comprises determining the first variation values according to

$$\mathrm{Adiff}(E_i, E_j) = (\Sigma_m \, | e_{i,m} - e_{j,m} |)/n$$

where Adiff is said first variation,

E_i is said first exponent set and E_j is a subsequent exponent set with $i > j$,

$E_i = (e_{i,0}, e_{i,1}, e_{i,2}, \ldots, e_{i,n-1})$,

$E_j = (e_{j,0}, e_{j,1}, e_{j,2}, \ldots, e_{j,n-1})$,

n is an integer representing the number of exponents in a said set of exponents,

$m = 0, 1, 2, \ldots, n-1$.

25. A method as claimed in claim 24, wherein the processing carried out by the weighted routing stage and first neural layer includes determining

$$Y = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \dots \\ \dots \\ y_{b-1} \end{bmatrix} = \Gamma\left(\begin{bmatrix} 0 & 0 & 0 & \dots & \dots & 0 \\ 0 & w_{1,1} & 0 & \dots & \dots & 0 \\ 0 & w_{2,1} & w_{2,2} & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & w_{b-1,1} & w_{b-1,2} & \dots & \dots & w_{b-1,b-1} \end{bmatrix} \begin{bmatrix} 0 \\ z_1 \\ z_2 \\ \dots \\ \dots \\ z_{b-1} \end{bmatrix} - \begin{bmatrix} 0 \\ T_1 \\ T_2 \\ \dots \\ \dots \\ T_{b-1} \end{bmatrix} \right)$$

the operator $\Gamma[\cdot]$ is defined as:

$$\Gamma \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \dots \\ \dots \\ \gamma_{b-1} \end{bmatrix} = \begin{bmatrix} f(\gamma_0) \\ f(\gamma_1) \\ f(\gamma_2) \\ \dots \\ \dots \\ f(\gamma_{b-1}) \end{bmatrix}$$

where $f(\gamma_i)$ is $+1$ if $\gamma \geq 0$ else it is 0,

Y represents outputs of the first neural layer,

T are threshold values determined during a training phase,

w are weighting values determined during the training phase, and

b is the number of exponent sets in the sequence comprising said data.

26. A method as claimed in claim 25, wherein the selection stage comprises selecting an output y_a of the first neural layer such that $y_a = 1$ and a is maximum for $i < a < b$.

27. A method as claimed in claim 26, wherein the plurality of exponent coding strategies comprises strategies $S_1, S_2, \dots, S_c$ where $c \leq b$, corresponding to respective differential coding limits $1, 2, \dots, c$.

28. A method as claimed in claim 27, wherein the exponent coding strategy S_v assigned to said first exponent set E_i is selected according to

$$\gamma = \max[\min(a+1, \sigma(E_i)), 1]$$

where $$\sigma(E_i) = \text{floor}((\textstyle\sum_j \| e_{i,j+1} - e_{i,j} \| / n) + 0.5).$$

29. A digital audio encoder in which audio data is transformed into coefficients having mantissas and exponents arranged in a sequence of sets, having:

a first variation processor coupled to receive the exponents of sets from said sequence and determine a first variation of exponent values between a first set and a plurality of subsequent sets in the sequence;

a second variation processor coupled to receive the exponents of said first set and determine a second variation between consecutive exponent values within said first set; and

a neural network processor coupled to receive said first and second variations and select and assign an exponent coding strategy to said first set from a plurality of coding strategies on the basis of said first and second variations, wherein each of the plurality of coding strategies correspond to a different differential coding limit.

30. An audio encoder as claimed in claim 29, wherein the neural network processor also selects and assigns an exponent coding strategy to at least one of the subsequent sets.

31. An audio encoder as claimed in claim 30, wherein the exponent coding strategy assigned to the at least one subsequent sets is an exponent re-use strategy.

32. An audio encoder as claimed in claim 29, wherein the neural network processor includes a weighted routing stage in which said first variation values are weighted according to predetermined weighting values and routed to inputs of a first neural layer, a selection stage in which an output of the first neural layer is selected, and an output processing stage in which a coding strategy is assigned to said first exponent set based on the output of said selection stage and said second variation.

33. An audio encoder as claimed in claim 32, wherein the first variation processor is arranged to determine said first variation according to

$$\mathrm{Adiff}(E_i, E_j) = (\Sigma_m |e_{i,m} - e_{j,m}|)/n$$

where Adiff is said first variation,

E_i is said first exponent set and E_j is a subsequent exponent set with $i > j$,

$E_i = (e_{i,0}, e_{i,1}, e_{i,2}, \ldots, e_{i,n-1})$,

$E_j = (e_{j,0}, e_{j,1}, e_{j,2}, \ldots, e_{j,n-1})$,

n is an integer representing the number of exponents in a said set of exponents,

$m = 0, 1, 2, \ldots, n-1$.

34. An audio coder as claimed in claim 29 or 33, wherein the weighted routing stage of the neural network processor is arranged to determine

$$Y = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \ldots \\ \ldots \\ y_{b-1} \end{bmatrix} = \Gamma\left(\begin{bmatrix} 0 & 0 & 0 & \ldots & \ldots & 0 \\ 0 & w_{1,1} & 0 & \ldots & \ldots & 0 \\ 0 & w_{2,1} & w_{2,2} & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & w_{b-1,1} & w_{b-1,2} & \ldots & \ldots & w_{b-1,b-1} \end{bmatrix} \begin{bmatrix} 0 \\ z_1 \\ z_2 \\ \ldots \\ \ldots \\ z_{b-1} \end{bmatrix} - \begin{bmatrix} 0 \\ T_1 \\ T_2 \\ \ldots \\ \ldots \\ T_{b-1} \end{bmatrix} \right)$$

the operator $\Gamma[\bullet]$ is defined as:

$$\Gamma \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \ldots \\ \ldots \\ \gamma_{b-1} \end{bmatrix} = \begin{bmatrix} f(\gamma_0) \\ f(\gamma_1) \\ f(\gamma_2) \\ \ldots \\ \ldots \\ f(\gamma_{b-1}) \end{bmatrix}$$

where $f(\gamma_i)$ is $+1$ if $\gamma \geq 0$ else it is 0.

Y represents outputs of the first neural layer,

T are threshold values determined during a training phase,

w are weighting values determined during the training phase, and

b is the number of sets in the sequence.

35. An audio encoder as claimed in claim 34, wherein the selection stage comprises selecting an output y_a of the first neural layer such that $y_a = 1$ and a is maximum for $i < a < b$.

36. A method as claimed in claim 35, wherein the plurality of exponent coding strategies comprises strategies $S_1, S_2, \ldots, S_c$, where $c \leq b$, corresponding to respective differential coding limits 1, 2, ... ,c.

37. A method as claimed in claim 36, wherein the exponent coding strategy S_γ assigned for encoding exponents in said first set E_i is selected according to

$$\gamma = \max[\min(a+1, \sigma(E_i)), 1]$$

where $$\sigma(E_i) = \mathrm{floor}((\textstyle\sum_j \|e_{i,j+1} - e_{i,j}\|/n) + 0.5).$$

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
H04B 1/66

A1

(11) International Publication Number: WO 99/41844

(43) International Publication Date: 19 August 1999 (19.08.99)

(21) International Application Number: PCT/SG98/00009

(22) International Filing Date: 12 February 1998 (12.02.98)

(71) Applicant (*for all designated States except US*): SGS-THOMSON MICROELECTRONICS ASIA PACIFIC (PTE) LTD. [SG/SG]; 28 Ang Mo Kio Industrial Park 2, Singapore 569508 (SG).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): ABSAR, Mohammed, Javed [IN/SG]; Block 411, #09–1006 Hougang Avenue 10, Singapore 530411 (SG). GEORGE, Sapna [IN/SG]; Block 315, #06–220 Serangoon Avenue 2, Singapore 550315 (SG). ALVAREZ-TINOCO, Antonio, Mario [GB/SG]; 32 Toh Tuck Road #03–01, Singapore 596710 (SG).

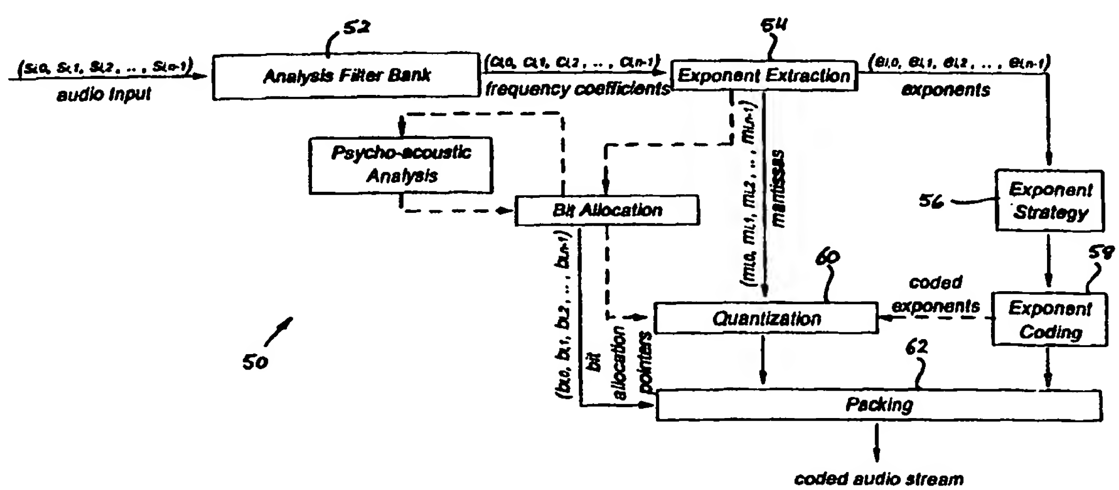
(74) Agent: DONALDSON & BURKINSHAW; P.O. Box 3667, Singapore 905667 (SG).

(81) Designated States: JP, SG, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: A NEURAL NETWORK BASED METHOD FOR EXPONENT CODING IN A TRANSFORM CODER FOR HIGH QUALITY AUDIO

AUDIO ENCODER



(57) Abstract

A method and apparatus for assigning an exponent coding strategy in a digital audio transform coder. Different coding strategies having different differential coding limits may be assigned to different set of transform exponents according to the frequency domain characteristics of the audio signal. A neural network processing system is utilised to perform an efficient mapping of each exponent set to an appropriate coding strategy.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

A NEURAL NETWORK BASED METHOD FOR EXPONENT CODING IN A TRANSFORM CODER FOR HIGH QUALITY AUDIO

Technical Field

This invention is applicable in the field of audio coders which employ exponent coding techniques to provide high levels of digital data compression.

Background Art

In order to more efficiently broadcast or record audio signals, the amount of information required to represent the audio signals may be reduced. In the case of digital audio signals, the amount of digital information needed to accurately reproduce the original pulse code modulation (PCM) samples may be reduced by applying a digital compression algorithm, resulting in a digitally compressed representation of the original signal. The goal of the digital compression algorithm is to produce a digital representation of an audio signal which, when decoded and reproduced, sounds the same as the original signal, while using a minimum of digital information for the compressed or encoded representation.

Recent advances in audio coding technology have led to high compression ratios while keeping audible degradation in the compressed signal to a minimum. These coders are intended for a variety of applications, including 5.1 channel film soundtracks, HDTV, laser discs and multimedia. Description of one applicable method can be found in the Advanced Television Systems Committee (ATSC) Standard document entitled "Digital Audio Compression (AC-3) Standard", Document A/52, 20 December, 1995.

In the basic approach, at the encoder the time domain audio signal is first converted to the frequency domain using a bank of filters. The frequency domain coefficients, thus generated, are converted to fixed point representation. In fixed point syntax, each coefficient is represented as a mantissa and an exponent. The bulk of the compressed

bitstream transmitted to the decoder comprises these exponents and mantissas.

The exponents are usually transmitted in their original form. However, each mantissa must be truncated to a fixed or variable number of decimal places. The number of bits to be used for coding each mantissa is obtained from a bit allocation algorithm which may be based on the masking property of the human auditory system. Lower numbers of bits result in higher compression ratios because less space is required to transmit the coefficients. However, this may cause high quantization errors, leading to audible distortion. A good distribution of available bits to each mantissa forms the core of the advanced audio coders.

Further compression is possible by employing differential coding for the exponents. In this case, the exponents for a channel are differentially coded across the frequency range. That is, instead of sending actual values, for many exponents only the difference between adjacent exponent values is sent. Furthermore, instead of coding every exponent a single exponent value is sent for every two exponents. In the extreme case, when exponent sets of several consecutive blocks in a frame are almost identical the exponent set for the first block is only sent, and the remaining blocks in the frame reuse the previously sent values.

These different methods amount to different exponent coding schemes presently in use, and are commonly referred to as exponent strategies.

The differential exponent coding schemes presently in use limit the difference between consecutive exponents to a fixed value which is predefined by the standard. This does not take into consideration the signal characteristic. As a result, sometimes for a fast varying audio spectrum the limit is too low, which results in information loss due to truncation. At other times the limiting value may be too large for a slow varying spectrum, thereby resulting in bit wastage.

A good set of exponent strategies is therefore desirable, together with an efficient

processing engine which determines which exponent strategy is most suited for a particular set of exponents. The processing engine should not be too computationally intensive to allow encoders to operate in real time and run on systems having relatively small computational resources.

Summary of the Invention

In accordance with the present invention, there is provided a method for coding audio data comprising a sequence of exponent sets each comprising a plurality of exponents, comprising the steps of:

determining a first variation of exponent values between a first exponent set in the sequence and each subsequent exponent set in said sequence;

selecting an exponent coding strategy for said first exponent set from a plurality of exponent coding strategies on the basis of said first variation; and

coding said first exponent set according to the selected exponent coding strategy.

Preferably each of the plurality of exponent coding strategies corresponds to a different differential coding limit.

In one form of the invention, the method includes selecting one of said subsequent exponent sets on the basis of said first variation and assigning an exponent re-use coding strategy to the selected exponent set and any exponent sets in said sequence between the first exponent set and the selected exponent set.

Preferably the method includes a step of determining a second variation between consecutive exponents in said first exponent set, wherein the exponent coding strategy for said first exponent set is selected on the basis of said first and second variations.

The step of selecting the exponent coding strategy for the first exponent set is preferably performed using neural network processing. In that case, the neural network processing

may include a feature extraction stage in which said sequence of exponent sets is processed to determine said first variation values, a weighted routing stage in which said first variation values are weighted according to predetermined weighting values and routed to inputs of a first neural layer, a selection stage in which an output of the first neural layer is selected, and an output processing stage in which a coding strategy is assigned to said first exponent set based on the output of said selection stage and said second variation.

Preferably the feature extraction stage comprises determining the first variation values according to

$$Adiff(\mathbf{E}_i, \mathbf{E}_j) = (\Sigma_m \, |e_{i,m} - e_{j,m}|)/n$$

where $Adiff$ is said first variation,

$\mathbf{E}_i$ is said first exponent set and $\mathbf{E}_j$ is a subsequent exponent set with $j > i$,

$\mathbf{E}_i = (e_{i,0}, e_{i,1}, e_{i,2}, \ldots, e_{i,n-1})$,

$\mathbf{E}_j = (e_{j,0}, e_{j,1}, e_{j,2}, \ldots, e_{j,n-1})$,

n is an integer representing the number of exponents in a said set of exponents,

$m = 0, 1, 2, \ldots, n\text{-}1$.

Preferably the processing carried out by the weighted routing stage and first neural layer includes determining

$$Y = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \cdots \\ \cdots \\ y_{b-1} \end{bmatrix} = \Gamma\left(\begin{bmatrix} 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & w_{1,1} & 0 & \cdots & \cdots & 0 \\ 0 & w_{2,1} & w_{2,2} & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & w_{b-1,1} & w_{b-1,2} & \cdots & \cdots & w_{b-1,b-1} \end{bmatrix} \begin{bmatrix} 0 \\ z_1 \\ z_2 \\ \cdots \\ \cdots \\ z_{b-1} \end{bmatrix} - \begin{bmatrix} 0 \\ T_1 \\ T_2 \\ \cdots \\ \cdots \\ T_{b-1} \end{bmatrix} \right)$$

the operator $\Gamma[\bullet]$ is defined as:

$$\Gamma\begin{bmatrix}\gamma_0\\ \gamma_1\\ \gamma_2\\ \cdots\\ \cdots\\ \gamma_{b-1}\end{bmatrix}=\begin{bmatrix}f(\gamma_0)\\ f(\gamma_1)\\ f(\gamma_2)\\ \cdots\\ \cdots\\ f(\gamma_{b-1})\end{bmatrix}$$

where $f(\gamma_i)$ is $+1$ if $\gamma \geq 0$ else it is 0,

Y represents outputs of the first neural layer,

T are threshold values determined during a training phase,

w are weighting values determined during the training phase, and

b is the number of exponent sets in the sequence comprising said data.

In one form of the invention, the selection stage comprises selecting an output y_a of the first neural layer such that $y_a = 1$ and a is maximum for $i<a<b$. The plurality of exponent coding strategies may comprise strategies $S_1, S_2, \ldots, S_c$, where $c \leq b$, corresponding to respective differential coding limits 1, 2, ... ,c, wherein the exponent coding strategy S_γ assigned to said first exponent set E_i is selected according to

$$\gamma = \max[\min(a+1, \sigma(E_i)), 1]$$

where $\sigma(E_i) = floor((\sum_j \| e_{i,j+1} - e_{i,j} \| / n) + 0.5)$.

The exponent sets in the sequence between $i+1$ and a, inclusive, can then be assigned the exponent re-use coding strategy.

The present invention also provides a method for processing data in an audio data encoder, the data comprising a sequence of exponent sets each comprising a plurality of exponents, comprising the steps of:

determining a first variation of exponent values within a first exponent set;

determining a second variation of exponent values between said first exponent set and each subsequent exponent set in said sequence; and

assigning an exponent coding strategy to the first exponent set based on the determined first and second variations.

The invention also provides a digital audio encoder in which audio data is transformed into coefficients having mantissas and exponents arranged in a sequence of sets, having:

a first variation processor coupled to receive the exponents of sets from said sequence and determine a first variation of exponent values between a first set and a plurality of subsequent sets in the sequence;

a second variation processor coupled to receive the exponents of said first set and determine a second variation between consecutive exponent values within said first set; and

a neural network processor coupled to receive said first and second variations and select and assign an exponent coding strategy to said first set from a plurality of coding strategies on the basis of said first and second variations.

Brief Description of the Drawings

The invention is described in greater detail hereinafter, by way of example only, with reference to several embodiments, and having regard to the accompanying drawings, wherein:

Figure 1 is a diagrammatic illustration of the data structure of an encoded audio data bit stream, having frames and blocks;

Figure 2 is a functional block diagram of an audio encoder;

Figure 3 is a diagrammatic illustration of mapping of exponent sets to coding strategies; and

Figure 4 is a diagrammatic illustration of a neural network system for selection of exponent coding strategies.

Detailed Description of the Preferred Embodiments

Suppose a frame comprises b blocks and a block $\mathbf{B}_i$ is composed of an ordered exponent set $\mathbf{E}_i$ $(e_{i,0}, e_{i,1}, e_{i,2}, \ldots, e_{i,n-1})$ and mantissa set $\mathbf{M}_i$ $(m_{i,0}, m_{i,1}, m_{i,2}, \ldots, m_{i,n-1})$. If the elements in $\mathbf{E}_i$ are very similar to those in $\mathbf{E}_{i+1}, \mathbf{E}_{i+2}, \ldots, \mathbf{E}_{i+a}$, then only exponent set $\mathbf{E}_i$ needs to be coded into the bitstream. Since the final question is more a matter of perception (the coding must not cause any audible degradation) rather than an exact closed form mathematical equation, a stochastic approach is suitable. Neural Networks are known to solve such problems efficiently with minimum computation requirements.

The maximum difference value allowed between exponents should be based on the variation in the exponents and the number of blocks over which the exponent set is to be reused. Allowing a higher value of maximum allowed difference results in less coding error. If an exponent value is to be reused several times it is necessary to code it with minimal error or else the accumulated error over several blocks could be quite appreciable.

Consider a set of exponent coding strategies $\{S_0, S_1, S_2, \ldots\}$. Let strategy S_0 be defined as *"reuse"*. That is, if exponent set $\mathbf{E}_i$, = $(e_{i,0}, e_{i,1}, e_{i,2}, \ldots, e_{i,n-1})$ uses strategy S_0 for coding then essentially no exponent is transmitted. Instead the receiver is directed to use exponents of $\mathbf{E}_{i-1}$ as those of exponent $\mathbf{E}_i$. Next, let S_1 be defined as an exponent coding strategy where all exponents are differentially encoded with the constraint that the difference between any two consecutive exponents is always -1, +1 or 0. That is, for exponent coding strategy S_1, the maximum allowed difference, **L**, equals +/-1. Similarly, for strategy S_2 , let the maximum allowed difference, **L**, be +/-2, and so on for $S_3, S_4, \ldots$.

The inputs $(\mathbf{E}_0, \mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_{b-1})$ are presented to the neural network system. The output $(o_0, o_1, o_2, \ldots, o_{b-1})$ of the system is the exponent strategy corresponding to each exponent set.

The input to an audio coder comprises a stream of digitised samples of the time domain analog signal, which is illustrated diagrammatically in Figure 1. The bit stream comprises a sequence of frames, with each frame consisting of b blocks identified as $(\mathbf{B}_0, \mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_{b-1})$. Each block contains n consecutive samples of data. A frame is an independent entity, such that there is no inter-frame data sharing. This facilitates splicing of encoded data at the frame level, and rapid recovery from transmission error. By grouping blocks to form a frame coding efficiency is increased since redundant data is shared, as permitted by signal conditions, across the frame. The block size in a particular coding standard is usually determined based on spectral and temporal resolution requirements for the audio signal.

Figure 2 illustrates a general block diagram of an audio encoder 50. Within any block $\mathbf{B}_i$ the n samples form an ordered set $(s_{i,0}, s_{i,1}, s_{i,2}, \ldots, s_{i,n-1})$. These time domain samples $(s_{i,0}, s_{i,1}, s_{i,2}, \ldots, s_{i,n-1})$ are converted to the frequency domain using an analysis filter bank 52. The frequency domain coefficients, thus generated, form another ordered set which can be identified as $(c_{i,0}, c_{i,1}, c_{i,2}, \ldots, c_{i,n-1})$. Here $c_{i,0}$ is the lowest frequency (dc) component while $c_{i,n-1}$ is the highest frequency component of the signal.

Audio compression essentially entails finding how much of the information in the set $(c_{i,0}, c_{i,1}, c_{i,2}, \ldots, c_{i,n-1})$ is necessary to reproduce the original analog signal at the decoder with minimal audible distortion.

The coefficient set $(c_{i,0}, c_{i,1}, c_{i,2}, \ldots, c_{i,n-1})$ is next converted into floating point format, where each coefficient $c_{i,j}$ is represented by an exponent $e_{i,j}$ and mantissa m_{ij}. By definition, $e_{i,j} = floor(log_2(c_{i,j}))$. Also, mantissa $m_{i,j} = (c_{i,j})/e_{i,j}$. The exponents and mantissas form two ordered sets $\mathbf{E}_i = (e_{i,0}, e_{i,1}, e_{i,2}, \ldots, e_{i,n-1})$ and $\mathbf{M}_i = (m_{i,0}, m_{i,1}, m_{i,2}, \ldots, m_{i,n-1})$. Therefore, each coefficient $c_{i,j}$ is related to the exponent $e_{i,j}$ and mantissa $m_{i,j}$ by the relation $c_{i,j} = m_{i,j} * 2^{e_{i,j}}$.

The exponent set is usually transmitted in its original form. However, the mantissa is truncated to a fixed or variable number of decimal places. Suppose mantissa $m_{i,j}$ is transmitted as $b_{i,j}$ number of bits. The value $b_{i,j}$ is usually obtained from a bit allocation algorithm which for advanced psychoacoustic coders may be based on the masking property of the human auditory system. Low value of $b_{i,j}$ results in higher compression ratio because less space is required to transmit the coefficients. However, this causes very high quantization error leading to audible distortion. A good distribution of available bits to each mantissa forms the core of the most advanced encoders.

Further compression is made possible in audio coders by application of certain coding techniques for exponents. Since most music is characterised by slowly changing spectrum across audio blocks, it is possible to reuse the exponents of a block $\mathbf{B}_i$ as exponents of the next few following blocks within the frame. That is, if the elements in exponent set $\mathbf{E}_i$ are very similar to the corresponding elements in sets $\mathbf{E}_{i+1}$, $\mathbf{E}_{i+2}$, ..., $\mathbf{E}_{i+a}$ then only exponent set $\mathbf{E}_i$ needs to be coded into the bitstream. A bit can be set to indicate to the decoder to *reuse* $\mathbf{E}_i$ for the next a blocks in the frame. The similarity of set $\mathbf{E}_i$ to another set $\mathbf{E}_j$ is, in general, a difficult problem to solve.

In addition to reuse, the exponents for a channel are usually differentially coded across the frequency range. That is, instead of sending the actual values, only the difference between the exponents are sent. However, as mentioned, the exponent coding schemes limit the allowable difference between adjacent exponents in the sequence to a fixed value, called **L**, which is predefined by the standard. This does not take into consideration the actual signal characteristics. As a result, sometimes for fast varying spectrum the limit is too low which results in information loss due to truncation. At other times the limiting value may be too large for a slow varying spectrum, thereby resulting in bit wastage.

Instead of allowing variation in value of **L**, some existing coders modify the frequency resolution in order to provide varying levels of compression for exponents. For example,

instead of sending delta values for every element in set $(e_{i,0}, e_{i,1}, e_{i,2}, e_{i,3}, e_{i,4} \ldots, e_{i,n-1})$, only delta for every 2 (or 3 or 4 etc.) frequency coefficients are sent. That is, the transmitted set is $(e_{i,0}, e_{i,2}, e_{i,4} \ldots, e_{i,n-1})$. At the receiver the values for $e_{i,1}, e_{i,3}, e_{i,5}$ are obtained by interpolation of the received values. However, this scheme is essentially a subset of the former scheme of varying the value of L.

To pose the problem more formally, suppose a set of exponent coding strategies is defined as $\{S_0, S_1, S_2, \ldots\}$. Let strategy S_0, be defined as "*reuse*". That is if exponent set $\mathbf{E}_i$ uses strategy S_0 for coding then essentially no exponent is transmitted but instead the receiver is directed to use exponents of $\mathbf{E}_{i-1}$ as those of exponent $\mathbf{E}_i$. Next, let S_1 be the exponent coding strategy where all exponents are differentially encoded with the constraint that the difference between any two consecutive exponents is always -1, +1 or 0. That is, for exponent coding strategy S_1, the maximum allowed difference, **L**, equals +/-1. Similarly, for strategy S_2, let the maximum allowed difference, **L**, be +/-2, and so on for $S_3, S_4, \ldots$.

Based on the above discussion, the problem of exponent strategy assignment can be formerly stated as a injective mapping of each exponent set $\mathbf{E}_i$ within a frame to a suitable strategy S_j such that maximum compression is achieved with minimal audible distortion. A mapping of exponent sets to coding strategies is illustrated diagrammatically in Figure 3. Since each frame is an independent entity, exponent set $\mathbf{E}_0$ is not allowed to use strategy S_0 or else it would attempt to use the exponent information in the previous frame.

For example, suppose a frame consists of six blocks with exponent sets $\mathbf{E}_0$, $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$, $\mathbf{E}_4$ and $\mathbf{E}_5$, and the assigned strategy is S_5, S_0 S_0, S_0, S_0 and S_1, respectively. This means exponents in set $\mathbf{E}_0$ are differentially encoded with the constraint that the maximum allowed difference is +/-5 (by previous definition of strategy S_5). For blocks 1, 2, 3 and 4, no exponent value is sent as they are instructed to reuse the exponents from the previously transmitted set. Block $\mathbf{B}_4$'s exponents, that is $\mathbf{E}_4$, are transmitted as

differentially encoded values with the constraint that the maximum allowed difference is +/-1 (definition of strategy S_1).

The optimal mapping is virtually impossible to find due to the complex behaviour of the system, large numbers of competing parameters, and the fact that optimality depends on perceived quality of reproduced music rather than on a closed form equation with a minimal value. Moreover, there is a long-standing debate on trade-offs between quality of final music, compression ratio achieved and computational requirements of the system. Therefore, having formally outlined the problem space, a stochastic approach is used to provide the solution. A Neural Network based solution is described herein.

Neural Network System

A neural network's ability to perform computation is based on the effort to reproduce the flexibility and power of the human brain by artificial means. Neural computation is performed by a dense mesh of computing nodes and connections. They operate collectively and simultaneously on most or all data inputs. The basic processing elements of the neural network are called nodes, or simply neurons. They can be considered as threshold units that fire when their total weighted input exceeds certain bias levels. Neurons are often organised in several layers, with neurons in consecutive layers often being connected together. Each connection strength is expressed by a numerical value called *weight*, the value of which is determined by the specific system which the network attempts to model.

Figure 4 is a diagrammatic illustration of a neural network system for selection of exponent coding strategies. The inputs to the neural network system are the exponent sets ($\mathbf{E}_0$, $\mathbf{E}_1$, $\mathbf{E}_2$, ..., $\mathbf{E}_{b-1}$), where $\mathbf{E}_i = (e_{i,0}, e_{i,1}, e_{i,2}, \ldots, e_{i,n-1})$. The output $(o_0, o_1, o_2, \ldots, o_{b-1})$ is the exponent strategy corresponding to each exponent set. Starting from $\mathbf{E}_0$ the strategy for each exponent set is determined. Suppose the exponent strategy for $\mathbf{E}_i$ is to be

determined. The exponent coding strategy for $\mathbf{E}_i$ depends on:

1. The number of times $\mathbf{E}_i$ will be reused across the blocks in the frame. If an exponent value is to be reused several times it is necessary to code it with minimal error or else the accumulated error over several blocks could be quite appreciable.

2. The variation between consecutive elements in the set $\mathbf{E}_i$. For a spectrum with large variations across the frequency range the limit $\mathbf{L}$ must be high to minimise information loss due to truncation. For a spectrum with low variation a low value of $\mathbf{L}$ is adequate and this results in better compression ratio.

A two stage network is used to take both these aspects into consideration when deciding on the strategy for an exponent set. The first criteria is of the number of blocks over which an exponent set is reused. This is dealt with primarily by the first layer. If the elements in $\mathbf{E}_i$ are very similar to those in $\mathbf{E}_{i+1}$, $\mathbf{E}_{i+2}$, ..., $\mathbf{E}_{1+a}$ then only exponent set $\mathbf{E}_i$ needs to be coded into the bitstream. A flag can be set to indicate to the decoder to *reuse* $\mathbf{E}_i$ for the next a blocks in the frame.

Feature Extraction

To find the number of blocks over which the same exponent set can be safely used without audible distortion in signal, a measuring function for similarity is required. The correlation coefficient:

$$r(E_i,E_j)=\frac{\sum_m (e_{i,m}-\overline{e}_i)*(e_{j,m}-\overline{e}_j)}{\sqrt{\sum_m (e_{i,m}-\overline{e}_i)^2*(e_{j,m}-\overline{e}_j)^2}} \qquad \text{Eq. 1}$$

is mathematically robust method for computing similarity. But it is too computationally

intensive for a practical system. Let *Adiff*($\mathbf{E}_i$, $\mathbf{E}_j$) be a function that determines the similarity between the elements of the set $\mathbf{E}_i$ and $\mathbf{E}_j$ and returns a value which shows the degree of difference. The similarity function *Adiff* (mean absolute difference) is defined as:

$$\mathrm{Adiff}(E_i, E_j) = (\Sigma_m \, |e_{i,m} - e_{j,m}|)/n \qquad \text{Eq. 2}$$

The values *Adiff*($\mathbf{E}_i$, $\mathbf{E}_{i+1}$), *Adiff*($\mathbf{E}_i$, $\mathbf{E}_{i+2}$), ..., *Adiff* ($\mathbf{E}_i$, $\mathbf{E}_{b-1}$) form the input $\mathbf{z}_1$, $\mathbf{z}_2$,... $\mathbf{z}_{b-1}$, respectively, to the first stage of neurons. Input $\mathbf{z}_0$ is always set to 0, as discussed hereinbelow. Since this step extracts only that characteristic of the input which is instrumental in the decision process of the network, it is termed as feature-extraction.

Reuse Strategy

The first stage of the network determines the maximum number of exponent sets that are considerably similar to a given set $\mathbf{E}_i$. The topology of the network and connection weights are such that the neuron $\mathbf{y}_1$ fires (corresponding output is 1) if $\mathbf{z}_1$ is within acceptable limit. Similarly, $\mathbf{y}_2$ fires if both $\mathbf{z}_1$ and $\mathbf{z}_2$ are within acceptable limits. Since $\mathbf{z}_1$ = *Adiff*($\mathbf{E}_i$, $\mathbf{E}_{i+1}$), it implies that if $\mathbf{y}_1$ fires then $\mathbf{E}_i$ and $\mathbf{E}_{i+1}$ are very similar. Similarly, if $\mathbf{y}_2$ fires it implies $\mathbf{E}_{i+1}$ and $\mathbf{E}_{i+2}$ are both very similar to $\mathbf{E}_i$. As $\mathbf{z}_0$ and the threshold value for $\mathbf{y}_0$ are both zero, $\mathbf{y}_0$ always fires.

Each neuron in the first layer therefore computes the weighted sum of its input. If the weighted sum is less than or equal to the threshold value, the neuron fires thus making the output 1. If the threshold value is larger than the weighted sum of inputs it means that not all inputs to this neuron are similar to E_1 and thus the neuron does not fire. Using matrix notation the output of the first stage of the network is:

$$Y = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \dots \\ \dots \\ y_{b-1} \end{bmatrix} = \Gamma\left(\begin{bmatrix} 0 & 0 & 0 & \dots & \dots & 0 \\ 0 & w_{1,1} & 0 & \dots & \dots & 0 \\ 0 & w_{2,1} & w_{2,2} & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & w_{b-1,1} & w_{b-1,2} & \dots & \dots & w_{b-1,b-1} \end{bmatrix} \begin{bmatrix} 0 \\ z_1 \\ z_2 \\ \dots \\ \dots \\ z_{b-1} \end{bmatrix} - \begin{bmatrix} 0 \\ T_1 \\ T_2 \\ \dots \\ \dots \\ T_{b-1} \end{bmatrix} \right) \qquad \text{Eq. 3}$$

the operator $\Gamma[\bullet]$ is defined as:

$$\Gamma \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \dots \\ \dots \\ \gamma_{b-1} \end{bmatrix} = \begin{bmatrix} f(\gamma_0) \\ f(\gamma_1) \\ f(\gamma_2) \\ \dots \\ \dots \\ f(\gamma_{b-1}) \end{bmatrix} \qquad \text{Eq. 4}$$

where $f(\gamma_i)$ is $+1$ if $\gamma \geq 0$ else it is 0. The weight $(w_{i,j})$ and the threshold (T_i) values are determined from the training phase of the neural network.

Perceptron learning rule is used for training the network. In this form of supervised training, the learning signal is a function of the difference between the desired and the actual neuron's response. Before the training begins, the weights are usually initialised to small random values (small values are necessary to avoid biased initial state). An input vector is presented to the system and the output is computed. The output from the network is compared with the desired output. The desired output was determined by extensive simulation and listening experiments. If the actual and the desired values are different the weights need to be adjusted to condition the network to imitate the desired behaviour. The weights and the threshold are adjusted according to Eq. 5 and Eq. 6, respectively. Here d_j is the desired output. The learning rate, $\propto = 0.11$, has been

experimentally found to be most suited for this application.

$$\Delta w_{j,k} = \alpha[d_j - f(\textstyle\sum_i (w_{j,i} * z_i) - T_j)] * z_k \qquad Eq.5$$

$$\Delta T_j = \alpha[d_j - f(\textstyle\sum_i (w_{j,i} * z_i) - T_j)] * 0.1 \qquad Eq.6$$

In this system, the threshold requires a multiplication by 0.1 for maintaining same rate of adjustments for both thresholds and weights. A good set of training vectors and well adjusted desired output are crucial to the correct setting of the weights.

An example of weights and thresholds of a trained system for a transform coder with six audio blocks is provided below:

$$W = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .05 & 0 & 0 & 0 & 0 \\ 0 & .91 & .23 & 0 & 0 & 0 \\ 0 & 1.2 & .96 & .29 & 0 & 0 \\ 0 & .96 & .92 & .59 & .20 & 0 \\ 0 & 1.2 & 1.0 & .94 & .88 & .28 \end{bmatrix} \qquad T = \begin{bmatrix} 0 \\ .02 \\ .33 \\ .71 \\ .75 \\ 1.2 \end{bmatrix}$$

Analysing the weight matrix W, an interesting behaviour becomes apparent. Along any row the weights decrease from left to right. The reason for this is as follows. Consider output $\mathbf{y}_2$.

$$y_2 = f(w_{2,1} * z_1 + w_{2,2} * z_2 - T_2)$$

$$= f(w_{2,1} * Adiff(E_i, E_{i+1}) + w_{2,2} * Adiff(E_i, E_{i+2}) - T_2) \qquad \text{Eq.7}$$

If we consider the spectrum of a signal across time, the spectrum shape at time t_{i+1} will be intermediate between spectrum shapes at time t_i and t_{i+2}, assuming $\Delta t = t_{i+1} - t_i = t_{i+2} - t_{i+1}$ is small enough. Therefore, if $\mathbf{E}_i$ reflects the spectral shape at time t_i and similarly

E_{i+1} and E_{i+2} the shape at t_{i+1} and t_{i+2}, respectively, then since the spectrum is changing over time, in general $Adiff(E_i, E_{i+1}) \leq Adiff(E_i, E_{i+2})$. In Eq. 7, to level this discrepancy the multiplicative weight for $Adiff(E_i, E_{i+1})$ is larger than that for $Adiff(E_i, E_{i+2})$.

In effect, the similarity function for estimating similarity of two exponent sets, $Adiff(E_i, E_j)$, now has a multiplicative factor $\vartheta(i,j) = w_{i,j}$. The new similarity function $\vartheta(i,j) * Adiff(E_i, E_j)$, judges similarity of two sets based on not only the mean absolute difference between the elements, but also the time difference expressed as difference in block number, that exists between the two exponents sets.

Two exponents sets closer in time are expected to match much more than two sets with much greater time difference, in order for the similarity function to generate the same value. The fact that the network was able to automatically detect this peculiar characteristic shows its strength and our correctness in modelling of the system. It is to be noted that the system is extremely robust. In an existing system new weight vectors, obtained from perhaps superior training data and training-rule, can always be updated to provide better results.

Winner Takes all Strategy

To the careful reader it is evident that at any one time it is possible that more than one neuron fires. The objective is to determine the maximum number of exponent sets which can be identified as similar to given set E_i. In the winner-takes-all strategy, starting from output y_{b-1} and proceeding upward, each neuron is analysed to find the first one that has fired. Suppose y_a is the first one from the bottom to have fired. Neuron y_a is declared winner and it's output is left as 1. All other neurons in this layer are now switched off. As a result the second layer's output now entirely dependent on neuron y_a.

From the network topology and the previous discussions, it is evident that y_0 always fires.

This ensures that there is always at least one winner. Now, neuron y_1 fires if exponent sets E_i and E_{i+1} are fond by the first layer to be very similar. Similarly, y_2 fires if exponent set E_{i+1} and E_{i+2} are both very similar to E_i.

In general, y_a fires when E_{i+1}, E_{i+2}, E_{i+3} ..., and E_{i+a} are all very similar to E_i. Therefore, exponent set E_i is only coded and transmitted into the bitstream. Exponent sets E_{i+1}, E_{i+2}, E_{i+3} ..., E_{i+a} all are forced to reuse the values in E_1 (that is, all are assigned strategy S_0).

The exact strategy for E_1 still needs to be determined from the available options $\{S_1, S_2, S_3 \ldots\}$. A simple technique would be to always use a fixed strategy for E_i (e.g. S_2) to provide minimal complexity.

A slight improvement on it would be to choose a strategy depending on the number of blocks over which the exponent set E_i would be reused. Note that if y_a is the winner then exactly a+1 blocks are using the same exponent sets. If an exponent value is to be reused several times it is meaningful to code it with minimal error or else the accumulated error over several blocks could be quite appreciable. Therefore one could choose strategy S_{a+1}, where y_a is the winner neuron. Thus by this process, if the winner neuron is y_a, the strategy for E_i, E_{i+1}, E_{i+2}, E_{i+3} ..., E_{i+a} is S_{a+1}, S_0, S_0, S_0, ..., S_0, respectively.

The problem with this improved version is that it does not take the variations of the elements within the exponent set into consideration. A much improved coding strategy for E_i would take into consideration not only the number of reuse blocks but also the variation within E_i.

Mean absolute deviation, defined as $Adev(E_i) = (\sum_j |e_{i,j} - \bar{e}_i|)/n)$ where $\bar{e}_i = \sum e_{i,j}/n$, could be used to provide a measure of variation. However, because finally differential coding is to be done, it is much more meaningful to base the strategy on the mean average difference

between consecutive exponents, rather than the deviation of exponents from the mean. Consequently, we define mean absolute deviation as:

$$\sigma(E_i) = floor((\sum_j \| e_{i,j+1} - e_{i,j} \| /n) + 0.5) \qquad \text{Eq. 8}$$

Finally, the exponent coding strategy recommended for $\mathbf{E}_i$ is $\mathbf{S}_\gamma$, where γ is defined by Eq. 9 below:

$$\gamma = \max[\min(a+1, \sigma(E_i)), 1] \qquad \text{Eq. 9}$$

By this definition, γ is always greater or equal 1. This ensures that reuse is never suggested for $\mathbf{E}_i$. If variation value $\sigma(E_i)$ is greater than the number of reuse blocks, the value a+1 dictates which strategy is finally chosen. This provides a safety measure by preventing explosion in bit usage for large varying signals with limited reuse. When variation is low, as guided by $\sigma(E_i)$, bit usage is optimised by using only as many bits for the differential limit as required and not being overwhelmed by the large number of reuse blocks.

At the end of this process the strategy for $\mathbf{E}_i$, $\mathbf{E}_{i+1}$, $\mathbf{E}_{i+2}$, $\mathbf{E}_{i+3}$, ..., $\mathbf{E}_{i+a}$, is $\mathbf{S}_\gamma$, $\mathbf{S}_0$, $\mathbf{S}_0$, $\mathbf{S}_0$, ..., $\mathbf{S}_0$, respectively. Now strategy for $\mathbf{E}_{i+a+1}$, $\mathbf{E}_{i+a+2}$, $\mathbf{E}_{i+a+3}$, ..., $\mathbf{E}_{b-1}$, needs to be determined. Treating $\mathbf{E}_{i+a+1}$ as $\mathbf{E}_i$ the whole process is repeated.

The deviation values $\sum_j \| e_{i,j+1} - e_{i,j} \| /n)$ for some benchmark music streams are given in the tables below. From the statistics it is evident that there is enough variation in frequency domain for commonly occurring music streams to warrant a flexible limit for L in a differential coding scheme. Moreover, from the range of values, it can be concluded that absolute deviation can be effectively used to determine the maximum differential limit, and hence the exponent strategy.

Adev	3.5	2.7	3.7	2.8	3.7	2.3	2.3
Reuse	2	1	3	3	2	4	3

Castanets

Adev	1.6	2.5	1.8	2.8	2.4	1.5	1.4
Reuse	6	6	6	4	6	6	6

Pop

Adev	4.2	3.13	3.03	2.6	2.5	2.7	0.9
Reuse	2	1	4	4	6	3	1

Harp

Note the third and fourth column in the table for castanets. The exponents are reused across three blocks in both cases but the variations are radically different. In the first case a differential limit of 4 is necessary while in the second a limit of 3 should be adequate. This shows that improvement in quality and coding efficiency can be achieved using this technique proposed above.

In summary of the above, referring again to the audio encoder shown in Figure 2, the frequency domain coefficients which are obtained from the analysis filter bank 52 are split into their constituent mantissas and exponents, and the exponents extracted (54) and passed to and exponent strategy processor 56. The exponent strategy processor 56 embodies the neural network system described hereinabove and illustrated diagrammatically in Figure 4. The exponent strategy processor 56 receives the extracted exponents and for each operation of the neural network system determines an appropriate coding strategy for at least one of the sets of exponents received. Each subsequent operation of the exponent strategy processor 56 then begins with the next exponent set not yet allocated a strategy. The exponents and assigned strategies are passed then to an exponent coder 58 which differentially codes the exponent sets according to the assigned strategies. The coded

exponents are passed to a quantizer 60 where they are quantized together with the corresponding mantissas for packing into the output bitstream. Where an exponent set has been allocated a re-use strategy, quantization is not required, and accordingly a "re-use flag" is passed directly from the exponent coder 58 to the bitstream packer 62.

The foregoing detailed description of the invention has been presented by way of example only, and is not intended to be considered limiting to the invention as defined in the claims appended hereto.

Claims:

1. A method for processing data in an audio data encoder, the data comprising a sequence of exponent sets each comprising a plurality of exponents, comprising the steps of:

determining a first variation of exponent values within a first exponent set;

determining a second variation of exponent values between said first exponent set and each subsequent exponent set in said sequence; and

assigning an exponent coding strategy to the first exponent set based on the determined first and second variations.

2. A method as claimed in claim 1, wherein the exponent coding strategy is assigned from a plurality of exponent coding strategies having different differential coding limits.

3. A method as claimed in claim 2, including a step of coding said first exponent set according to the assigned exponent coding strategy.

4. A method as claimed in claim 3, including a step of assigning an exponent coding strategies to at least one subsequent exponent set based on the corresponding determined second variation.

5. A method as claimed in claim 4, wherein the plurality of exponent coding strategies includes an exponent set re-use strategy which is assigned to the at least one subsequent exponent set.

6. A method as claimed in claim 5, including a step of coding said first exponent set and said at least one subsequent exponent set according to the corresponding assigned coding strategies.

7. A method as claimed in claim 2 or 6, wherein the steps of determining the first and second variations are performed utilising neural network processing.

8. A method as claimed in claim 7, wherein the neural network processing includes first and second neural layers.

9. A method as claimed in claim 7, wherein the neural network processing comprises a feature extraction stage in which said sequence of exponent sets is utilised to determine said second variations, a weighted routing stage in which said second variations are weighted according to predetermined weighting values and routed to inputs of a first neural layer, a selection stage in which an output of the first neural layer is selected, and an output processing stage in which a coding strategy is assigned to said first exponent set based on said first variation and the output of said selection stage.

10. A method as claimed in claim 9, wherein a coding strategy is assigned to at least one subsequent exponent set on the basis of the output of said selection stage.

11. A method as claimed in claim 10, wherein the coding strategy assigned to the at least one subsequent exponent set is an exponent re-use strategy.

12. A method as claimed in claim 9, wherein the feature extraction stage comprises determining

$$\mathrm{Adiff}(E_i, E_j) = (\Sigma_m |e_{i,m} - e_{j,m}|)/n$$

where Adiff is said second variation,

E_i is said first exponent set and E_j is a subsequent exponent set with $j > i$,

$E_i = (e_{i,0}, e_{i,1}, e_{i,2}, \ldots, e_{i,n-1})$,

$E_j = (e_{j,0}, e_{j,1}, e_{j,2}, \ldots, e_{j,n-1})$,

n is an integer representing the number of exponents in a said set of exponents,

m = 0, 1, 2, ... , n-1.

13. A method as claimed in claim 12, wherein the processing carried out by the weighted routing stage and first neural layer includes determining

$$Y = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \cdots \\ \cdots \\ y_{b-1} \end{bmatrix} = \Gamma\left(\begin{bmatrix} 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & w_{1,1} & 0 & \cdots & \cdots & 0 \\ 0 & w_{2,1} & w_{2,2} & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & w_{b-1,1} & w_{b-1,2} & \cdots & \cdots & w_{b-1,b-1} \end{bmatrix} \begin{bmatrix} 0 \\ z_1 \\ z_2 \\ \cdots \\ \cdots \\ z_{b-1} \end{bmatrix} - \begin{bmatrix} 0 \\ T_1 \\ T_2 \\ \cdots \\ \cdots \\ T_{b-1} \end{bmatrix} \right)$$

the operator $\Gamma[\bullet]$ is defined as:

$$\Gamma \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \cdots \\ \cdots \\ \gamma_{b-1} \end{bmatrix} = \begin{bmatrix} f(\gamma_0) \\ f(\gamma_1) \\ f(\gamma_2) \\ \cdots \\ \cdots \\ f(\gamma_{b-1}) \end{bmatrix}$$

where $f(\gamma_i)$ is +1 if $\gamma \geq 0$ else it is 0,
Y represents outputs of the first neural layer,
T are threshold values determined during a training phase,
w are weighting values determined during the training phase, and
b is the number of exponent sets in the sequence comprising said data.

14. A method as claimed in claim 13, wherein the selection stage comprises selecting an output y_a of the first neural layer such that $y_a = 1$ and a is maximum for $i < a < b$.

15. A method as claimed in claim 14, wherein the plurality of exponent coding strategies comprises strategies S_1, S_2, ... , S_c, where $c \leq b$, corresponding to differential coding limits 1, 2, ... ,c.

16. A method as claimed in claim 15, wherein the exponent coding strategy S_γ assigned to said first exponent set E_i is selected according to

$$\gamma = \max[\min(a+1, \sigma(E_i)), 1]$$

where $$\sigma(E_i) = floor((\textstyle\sum_j \| e_{i,j+1} - e_{i,j} \| / n) + 0.5).$$

17. A method for coding audio data comprising a sequence of exponent sets each comprising a plurality of exponents, comprising the steps of:

determining a first variation of exponent values between a first exponent set in the sequence and each subsequent exponent set in said sequence;

selecting an exponent coding strategy for said first exponent set from a plurality of exponent coding strategies on the basis of said first variation; and

coding said first exponent set according to the selected exponent coding strategy.

18. A method as claimed in claim 17, wherein each of the plurality of exponent coding strategies corresponds to a different differential coding limit.

19. A method as claimed in claim 17, including selecting one of said subsequent exponent sets on the basis of said first variation and assigning an exponent re-use coding strategy to the selected exponent set and any exponent sets in said sequence between the first exponent set and the selected exponent set.

20. A method as claimed in claim 17, 18 or 19, including a step of determining a second variation between consecutive exponents in said first exponent set, wherein the exponent coding strategy for said first exponent set is selected on the basis of said first and second variations.

21. A method as claimed in claim 17, wherein the step of selecting the exponent coding strategy for said first exponent set is performed utilising neural network processing.

22. A method as claimed in claim 20, wherein the step of selecting the exponent coding strategy for said first exponent set is performed utilising neural network processing.

23. A method as claimed in claim 22, wherein the neural network processing comprises a feature extraction stage in which said sequence of exponent sets is processed to determine said first variation values, a weighted routing stage in which said first variation values are weighted according to predetermined weighting values and routed to inputs of a first neural layer, a selection stage in which an output of the first neural layer is selected, and an output processing stage in which a coding strategy is assigned to said first exponent set based on the output of said selection stage and said second variation.

24. A method as claimed in claim 23, wherein a coding strategy is assigned to at least one subsequent exponent set on the basis of the output of said selection stage.

25. A method as claimed in claim 24, wherein the coding strategy assigned to the at least one subsequent exponent set is an exponent re-use strategy.

26. A method as claimed in claim 23, wherein the feature extraction stage comprises determining the first variation values according to

$$Adiff(E_i, E_j) = (\Sigma_m |e_{i,m} - e_{j,m}|)/n$$

where Adiff is said first variation,

E_i is said first exponent set and E_j is a subsequent exponent set with $i > j$,

$E_i = (e_{i,0}, e_{i,1}, e_{i,2}, \ldots, e_{i,n-1})$,

$E_j = (e_{j,0}, e_{j,1}, e_{j,2}, \ldots, e_{j,n-1})$,

n is an integer representing the number of exponents in a said set of exponents,

m = 0, 1, 2, ... , n-1.

27. A method as claimed in claim 26, wherein the processing carried out by the weighted routing stage and first neural layer includes determining

$$Y = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \cdots \\ \cdots \\ y_{b-1} \end{bmatrix} = \Gamma\left(\begin{bmatrix} 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & w_{1,1} & 0 & \cdots & \cdots & 0 \\ 0 & w_{2,1} & w_{2,2} & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & w_{b-1,1} & w_{b-1,2} & \cdots & \cdots & w_{b-1,b-1} \end{bmatrix} \begin{bmatrix} 0 \\ z_1 \\ z_2 \\ \cdots \\ \cdots \\ z_{b-1} \end{bmatrix} - \begin{bmatrix} 0 \\ T_1 \\ T_2 \\ \cdots \\ \cdots \\ T_{b-1} \end{bmatrix} \right)$$

the operator $\Gamma[\bullet]$ is defined as:

$$\Gamma \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \cdots \\ \cdots \\ \gamma_{b-1} \end{bmatrix} = \begin{bmatrix} f(\gamma_0) \\ f(\gamma_1) \\ f(\gamma_2) \\ \cdots \\ \cdots \\ f(\gamma_{b-1}) \end{bmatrix}$$

where $f(\gamma_i)$ is +1 if $\gamma \geq 0$ else it is 0,

Y represents outputs of the first neural layer,

T are threshold values determined during a training phase,

w are weighting values determined during the training phase, and

b is the number of exponent sets in the sequence comprising said data.

28. A method as claimed in claim 27, wherein the selection stage comprises selecting

an output y_a of the first neural layer such that $y_a = 1$ and a is maximum for $i < a < b$.

29. A method as claimed in claim 28, wherein the plurality of exponent coding strategies comprises strategies $S_1, S_2, \ldots, S_c$, where $c \leq b$, corresponding to respective differential coding limits 1, 2, ... ,c.

30. A method as claimed in claim 29, wherein the exponent coding strategy S_γ assigned to said first exponent set E_i is selected according to

$$\gamma = \max[\min(a+1, \sigma(E_i)), 1]$$

where $\sigma(E_i) = floor((\sum_j \| e_{i,j+1} - e_{i,j} \| / n) + 0.5)$.

31. A digital audio encoder in which audio data is transformed into coefficients having mantissas and exponents arranged in a sequence of sets, having:

a first variation processor coupled to receive the exponents of sets from said sequence and determine a first variation of exponent values between a first set and a plurality of subsequent sets in the sequence;

a second variation processor coupled to receive the exponents of said first set and determine a second variation between consecutive exponent values within said first set; and

a neural network processor coupled to receive said first and second variations and select and assign an exponent coding strategy to said first set from a plurality of coding strategies on the basis of said first and second variations.

32. An audio encoder as claimed in claim 31, wherein each of the plurality of coding strategies correspond to a different differential coding limit.

33. An audio encoder as claimed in claim 31 or 32, wherein the neural network processor also selects and assigns an exponent coding strategy to at least one of the subsequent sets.

34. An audio encoder as claimed in claim 33, wherein the exponent coding strategy assigned to the at least one subsequent sets is an exponent re-use strategy.

35. An audio encoder as claimed in claim 31, wherein the neural network processor includes a weighted routing stage in which said first variation values are weighted according to predetermined weighting values and routed to inputs of a first neural layer, a selection stage in which an output of the first neural layer is selected, and an output processing stage in which a coding strategy is assigned to said first exponent set based on the output of said selection stage and said second variation.

36. An audio encoder as claimed in claim 35, wherein the first variation processor is arranged to determine said first variation according to

$$\mathrm{Adiff}(E_i, E_j) = (\Sigma_m |e_{i,m} - e_{j,m}|)/n$$

where Adiff is said first variation,

E_i is said first exponent set and E_j is a subsequent exponent set with $i > j$,

$E_i = (e_{i,0}, e_{i,1}, e_{i,2}, \ldots, e_{i,n-1})$,

$E_j = (e_{j,0}, e_{j,1}, e_{j,2}, \ldots, e_{j,n-1})$,

n is an integer representing the number of exponents in a said set of exponents,

$m = 0, 1, 2, \ldots, n-1$.

37. An audio coder as claimed in claim 31 or 36, wherein the weighted routing stage of the neural network processor is arranged to determine

$$Y = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \cdots \\ \cdots \\ y_{b-1} \end{bmatrix} = \Gamma\left(\begin{bmatrix} 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & w_{1,1} & 0 & \cdots & \cdots & 0 \\ 0 & w_{2,1} & w_{2,2} & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & w_{b-1,1} & w_{b-1,2} & \cdots & \cdots & w_{b-1,b-1} \end{bmatrix} \begin{bmatrix} 0 \\ z_1 \\ z_2 \\ \cdots \\ \cdots \\ z_{b-1} \end{bmatrix} - \begin{bmatrix} 0 \\ T_1 \\ T_2 \\ \cdots \\ \cdots \\ T_{b-1} \end{bmatrix} \right)$$

the operator $\Gamma[\bullet]$ is defined as:

$$\Gamma\begin{bmatrix}\gamma_0\\ \gamma_1\\ \gamma_2\\ \dots\\ \dots\\ \gamma_{b-1}\end{bmatrix} = \begin{bmatrix}f(\gamma_0)\\ f(\gamma_1)\\ f(\gamma_2)\\ \dots\\ \dots\\ f(\gamma_{b-1})\end{bmatrix}$$

where $f(\gamma_i)$ is $+1$ if $\gamma \geq 0$ else it is 0,

Y represents outputs of the first neural layer,

T are threshold values determined during a training phase,

w are weighting values determined during the training phase, and

b is the number of sets in the sequence.

38. An audio encoder as claimed in claim 37, wherein the selection stage comprises selecting an output y_a of the first neural layer such that $y_a = 1$ and a is maximum for $i<a<b$.

39. A method as claimed in claim 38, wherein the plurality of exponent coding strategies comprises strategies $S_1, S_2, \dots, S_c$, where $c \leq b$, corresponding to respective differential coding limits 1, 2, ... ,c.

40. A method as claimed in claim 39, wherein the exponent coding strategy S_γ assigned for encoding exponents in said first set E_i is selected according to

$$\gamma = \max[\min(a+1, \sigma(E_i)), 1]$$

where $\sigma(E_i) = floor((\sum_j |e_{i,j+1} - e_{i,j}|/n) + 0.5)$.

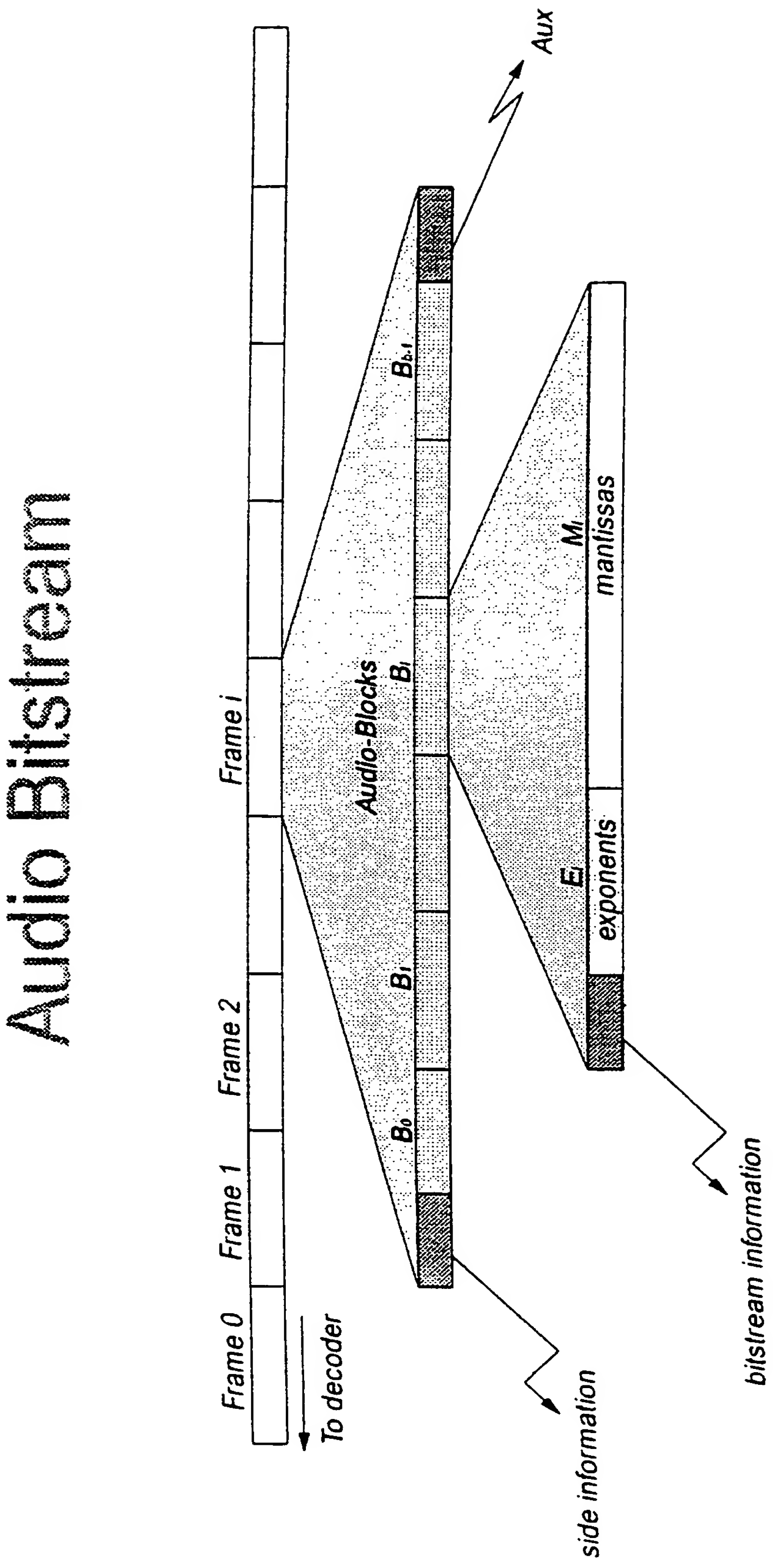

Audio Bitstream
Frame 0
Frame 1
Frame 2
Frame i
To decoder
Audio-Blocks
B_0
B_1
B_i
B_{b-1}
Aux
side information
E_i
exponents
M_i
mantissas
bitstream information

Fig. 1

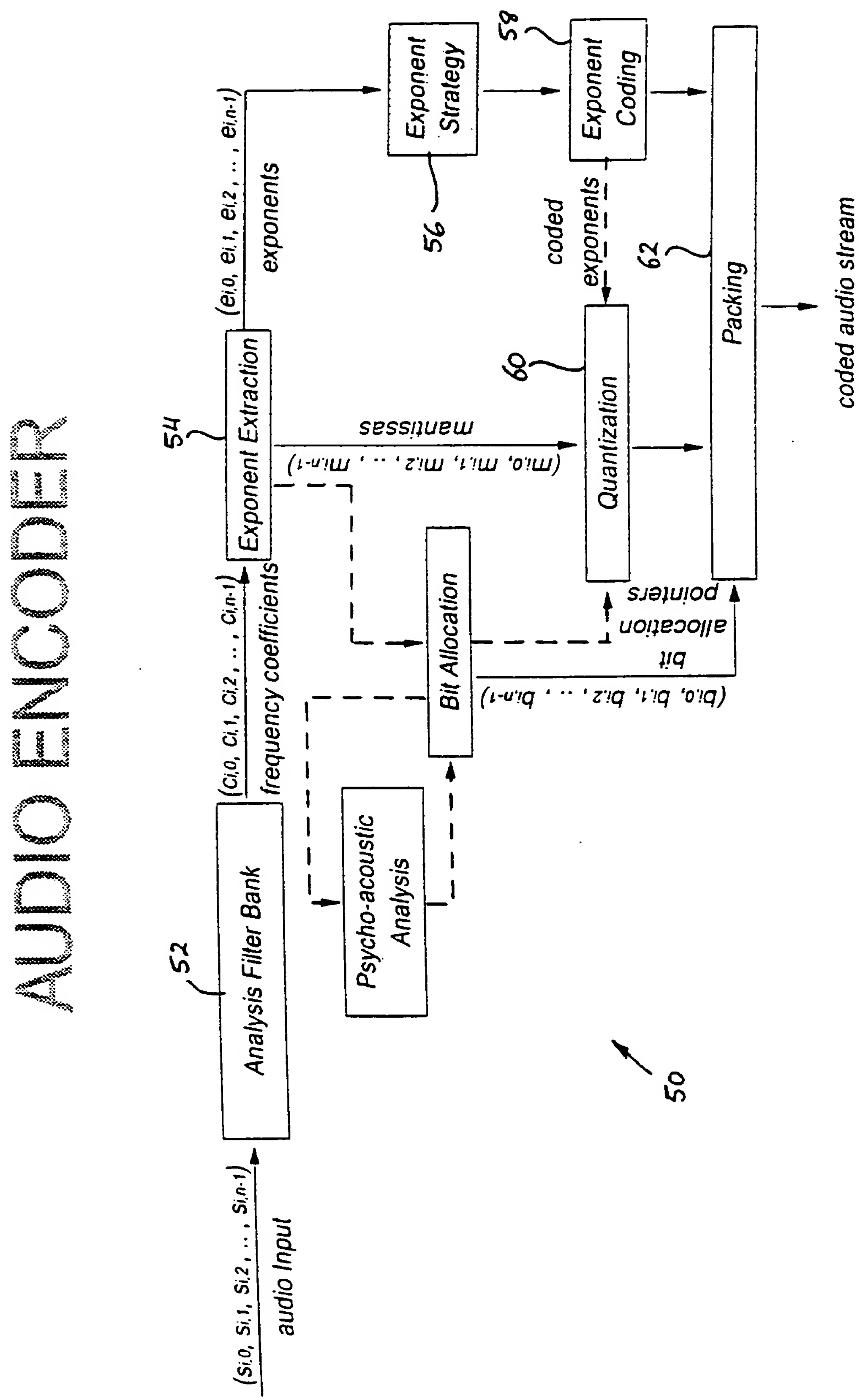

AUDIO ENCODER
$(s_{i,0}, s_{i,1}, s_{i,2}, \ldots, s_{i,n-1})$
audio Input
52
Analysis Filter Bank
$(c_{i,0}, c_{i,1}, c_{i,2}, \ldots, c_{i,n-1})$
frequency coefficients
Psycho-acoustic Analysis
54
Exponent Extraction
$(e_{i,0}, e_{i,1}, e_{i,2}, \ldots, e_{i,n-1})$
exponents
Exponent Strategy
56
58
Exponent Coding
coded exponents
mantissas
$(m_{i,0}, m_{i,1}, m_{i,2}, \ldots, m_{i,n-1})$
Bit Allocation
60
Quantization
bit allocation pointers
$(b_{i,0}, b_{i,1}, b_{i,2}, \ldots, b_{i,n-1})$
62
Packing
coded audio stream
50

Fig. 2

MAPPING EXPONENT SETS TO CODING STRATEGY

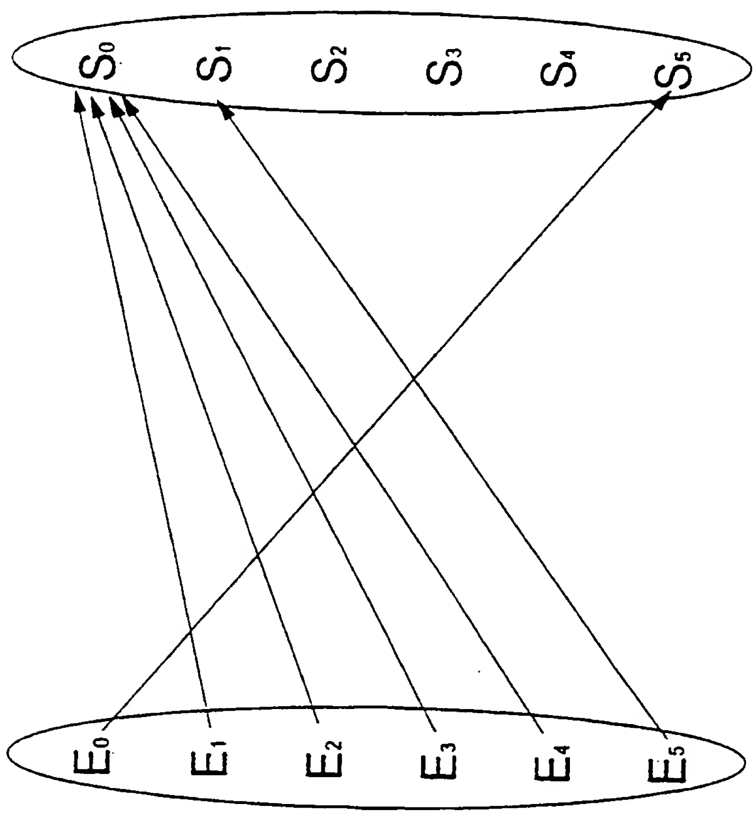


Fig. 3

Neural Network System For Exponent Coding Strategy

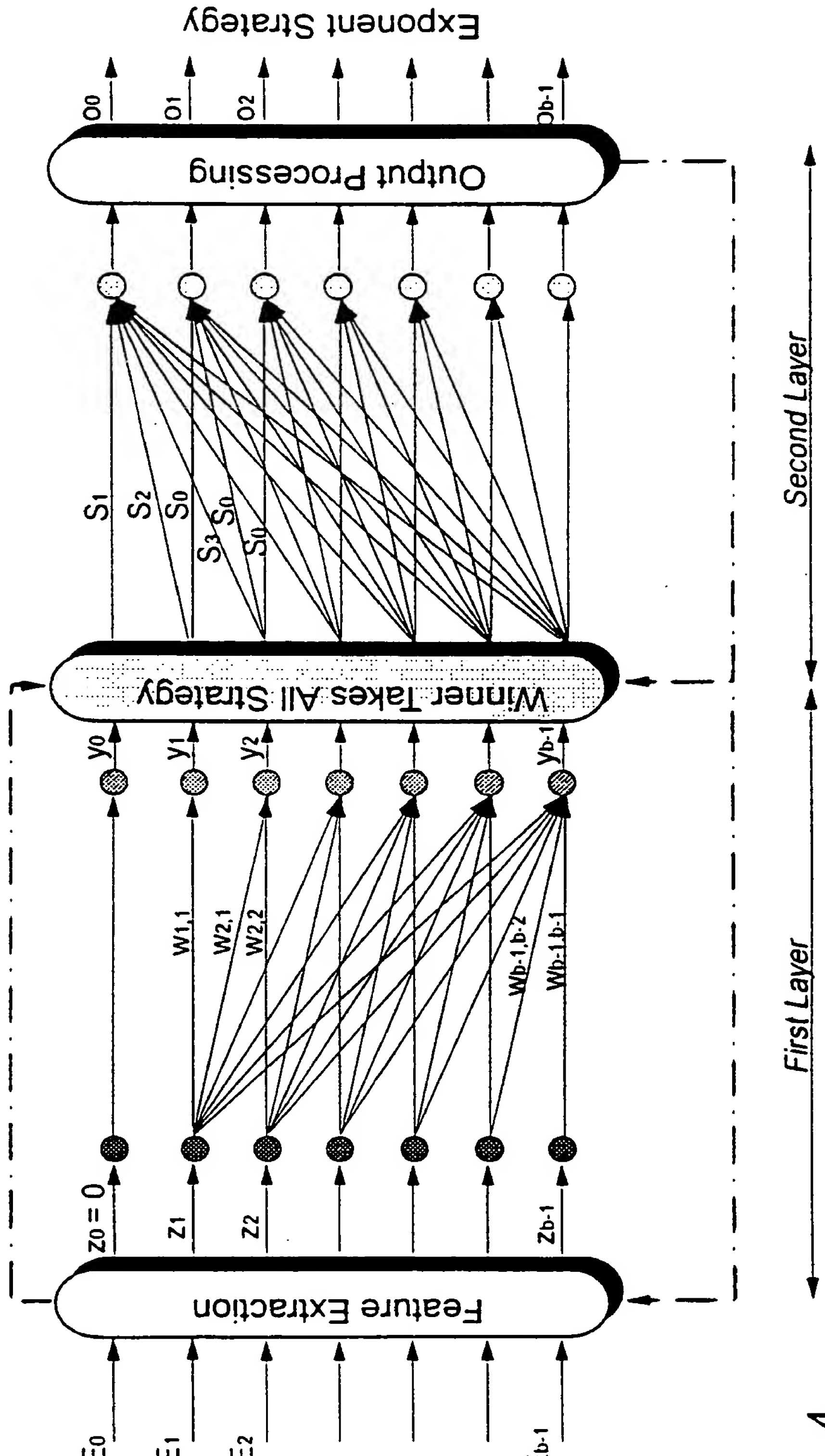


Fig. 4

INTERNATIONAL SEARCH REPORT

Int[illegible]al Application No
P[illegible]SG 98/00009

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 H04B1/66

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 H04B H03M

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MCKINNEY: "Digital Audio Compression Standard AC-3" 20 December 1995 , ADVANCED TELEVISION SYSTEMS COMITTEE XP002075746 cited in the application see paragraph 7.1.1 see paragraph 7.1.2	1,3-6, 17,19
A	see paragraph 8.2.8 -----	31

☐ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 October 1998	19/10/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Bossen, M

1